

ECE 517: MACHINE LEARNING

ASSIGNMENT 5: CONSTRUCTION OF A REGRESSION ALGORITHM

DAVID KIRBY – 101652098 – DAVIDKIRBY@UNM.EDU

FALL 2021



The .mat and the .csv files provided is a dataset that contains a training and a test sets. Each set consists of the predictor input X_{train} and X_{test} for regressors y_{train} and y_{test} , corresponding to the response y of a physical model to a vector signal x of 19 dimensions. The train and test outputs consist of 81 samples and they are depicted in the figure below.

In this assignment, you must construct a linear ridge regression that uses inputs X to predict y . Use 20% of the training data to validate gamma. The validation procedure consists of:

- Training the predictor with 80% of the training data and a given value for gamma.
- Running a test for the rest of the training data.
- Computing the mean square error of the prediction.
- Repeat for a reasonable range of gamma.

Choose the value of gamma that produced the best result. A reasonable interval can be between 0.01 times and 10 times the trace of the matrix, using a logarithmic spacing. In Matlab, use function logspace.

Provide the following results of the experiment:

- A graph comparing the real and predicted data.
- A graph of the result of the validation square error.
- The value of the optimal validation and test square errors.

Construction of a regression algorithm with the model parameters.

Using the example regression algorithm provided in the assignment, I swept gamma from 0.1 to 10 times the trace matrix.

The optimal value for γ was determined to be 0.004731 which led to an MMSE of 0.00808.

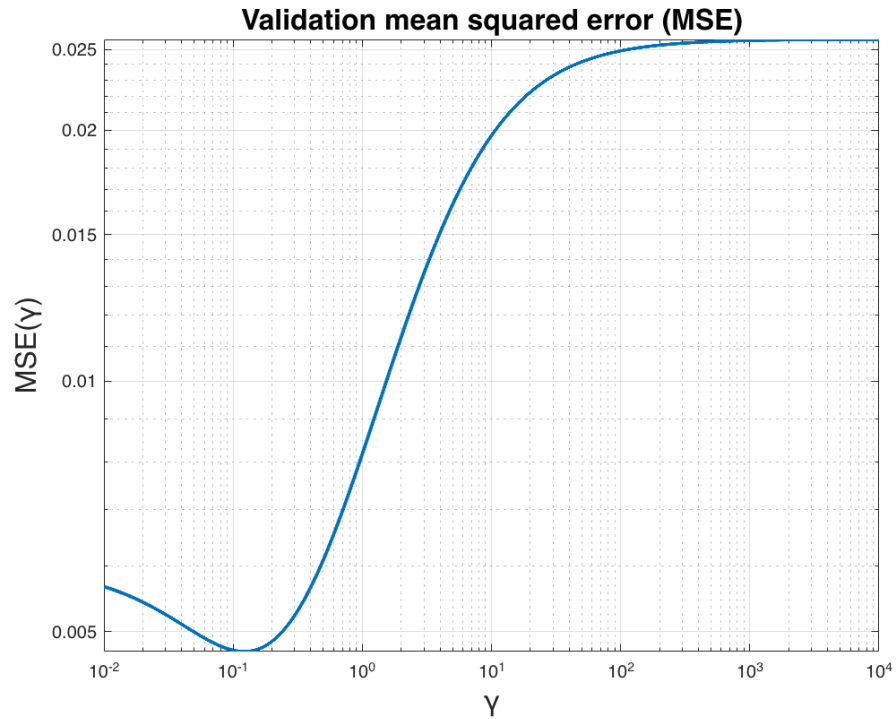


Figure 1: Validation mean square error (γ swept from 0.1 to 10).



Figure 2: Regression test results.

Repeat the experiment above, but using an SVR and a nu-SVR. In both cases you should use a 20% of the training data in order to validate the parameters of the SVM.

Provide the following results of the experiment:

- A graph comparing the real and predicted data.
- A graph of the result of the validation square error.
- The value of the optimal validation and test square errors.
- A written comparison of the results of this experiment and the previous ones.

Repeating the experiment using SVR and nu-SVR.

The optimal value for ϵ was determined to be 0.01 and $C = 5$ which led to an MMSE of 0.00576.

The optimal value for ν was determined to be 0.4 and $C = 10$ which led to an MMSE of 0.00557.

All three results are very similar in their results for the MMSE. This makes sense as we are testing all of these algorithms on the same dataset. The reason as to why this might be can be due to the values chosen for the cost and ϵ chosen for the SVR machine when creating the classifier/predictor. Similarly, the cost and ν values chosen for the nu-SVR machine could also be the reason for the slight difference in accuracy between the machines.

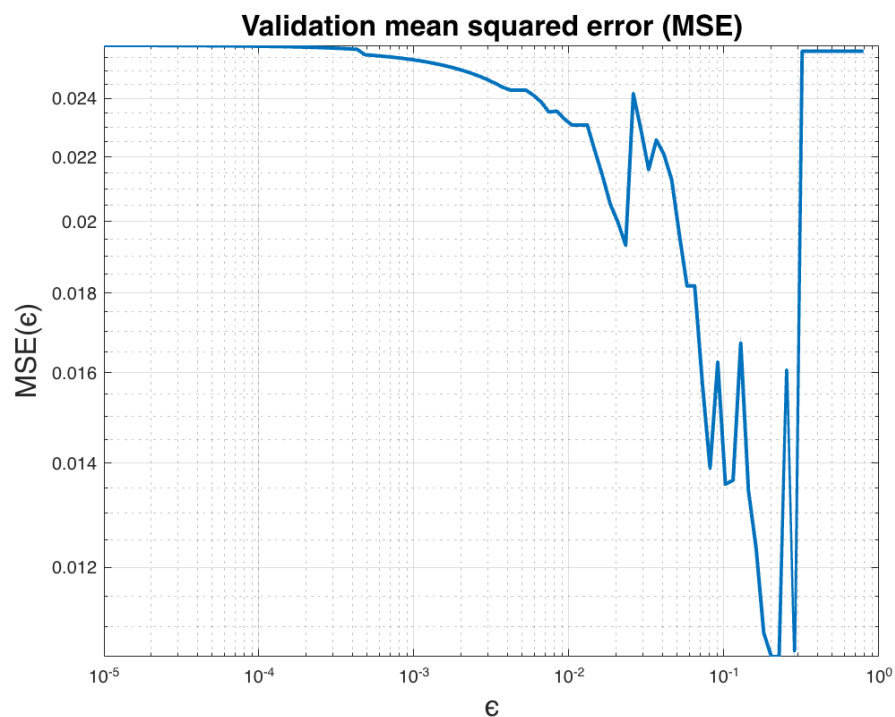


Figure 3: Validation mean square error (ϵ).

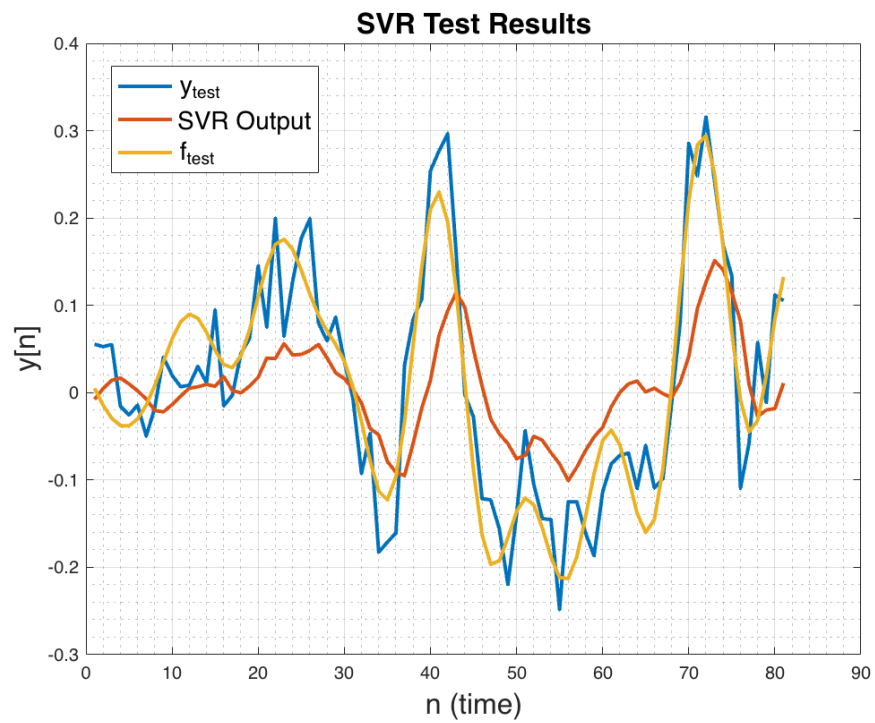
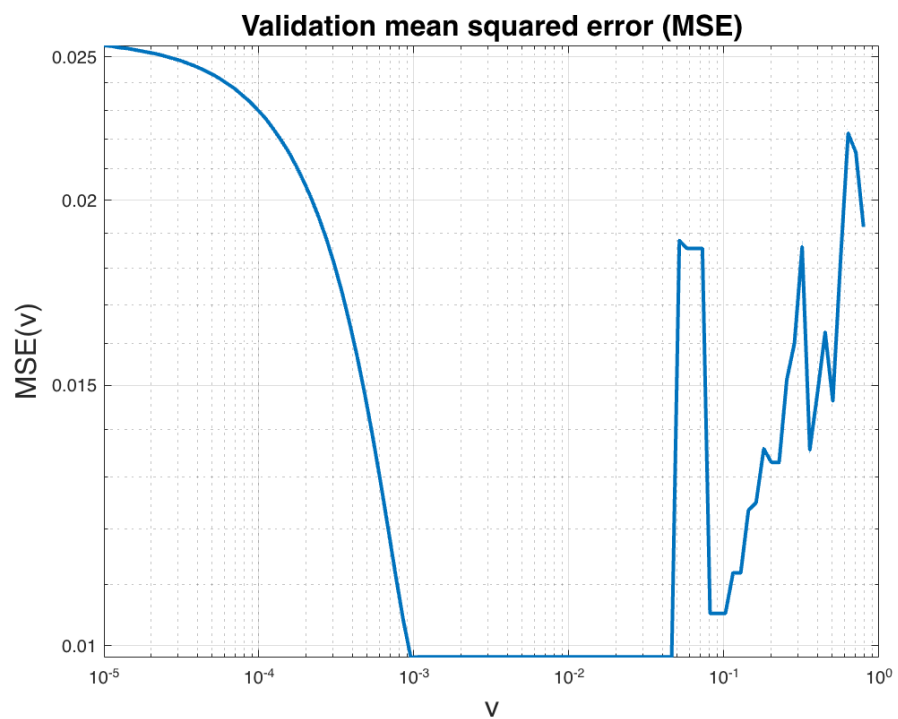


Figure 4: SVR test results.

Figure 5: Validation mean square error (ν).

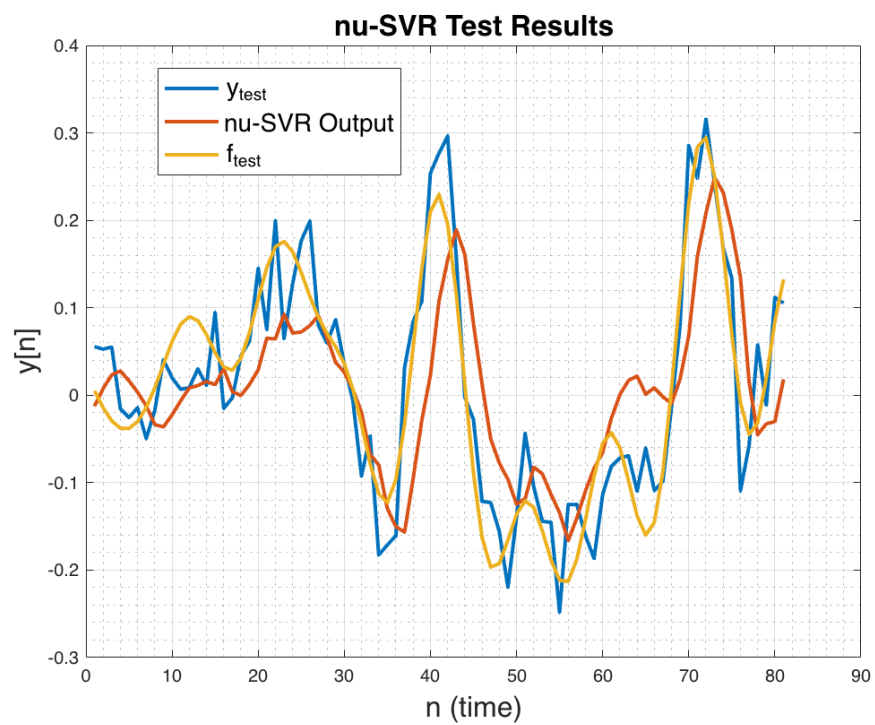


Figure 6: nu-SVR test results.